



Perceptually Relevant and Piecewise Linear Matching of Silhouettes

Zabulis, Xenophon; Sporring, Jon; Orphanoudakis, Xenophon

Published in:
Pattern Recognition

DOI:
[doi:10.1016/j.patcog.2004.06.003](https://doi.org/10.1016/j.patcog.2004.06.003)

Publication date:
2005

Document version
Early version, also known as pre-print

Citation for published version (APA):
Zabulis, X., Sporring, J., & Orphanoudakis, X. (2005). Perceptually Relevant and Piecewise Linear Matching of Silhouettes. *Pattern Recognition*, 38(1), 75-93. <https://doi.org/doi:10.1016/j.patcog.2004.06.003>



PERGAMON

Available at
www.ElsevierComputerScience.com

POWERED BY SCIENCE @ DIRECT®

Pattern Recognition 38 (2005) 75–93

PATTERN
RECOGNITION

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Perceptually relevant and piecewise linear matching of silhouettes

Xenophon Zabulis^{a,1}, Jon Sporring^{b,1}, Stelios C. Orphanoudakis^{c,*}

^aGRASP Laboratory, Levine Hall L402, University of Pennsylvania, 3330 Walnut Street, Philadelphia, PA 19104-6228, USA

^bDepartment of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen, Denmark

^cInstitute of Computer Science, FORTH, P.O. Box 1385, GR-71110 Heraklion, Crete, Greece

Received 2 December 2003; received in revised form 16 June 2004; accepted 16 June 2004

Abstract

In this paper, a novel alignment method for silhouettes is proposed. This method is based on the establishment of correspondences between landmarks on their boundaries and, in turn, on the establishment of correspondences of the boundary pieces in between these landmarks. The method yields more correct correspondences than conventional methods that scale the arc-length descriptions of silhouettes to align them. The selection of landmarks is investigated as to the robustness of their localization and their perceptual relevance. Matching of silhouettes is then achieved by quantifying the dissimilarity of a pair of silhouette boundaries, based on a novel dissimilarity metric. The matching procedure is evaluated, based on retrieval experiments, and it is concluded that the precision of the results is higher than that obtained by conventional pointwise comparison methods. © 2004 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

Keywords: Silhouette matching; Scale-space; Shape alignment; Curvature; Landmarks; Perception

1. Introduction

A significant component of image content, as appreciated by observers, is the shape of image contours. The ability to identify and retrieve similar contours is important in a wide spectrum of applications requiring content-based image retrieval and object recognition. A research challenge in any effort at automating this similarity matching process is how to make it perceptually relevant by emulating the corresponding behavior of human observers. It is expected that meeting this challenge would also contribute toward a more intuitive appreciation of content-based image retrieval (CBIR) results by end-users.

In this paper, a silhouette is considered as “the outline of a body viewed as circumscribing a mass” [1] and, thus, is considered without holes. The paper focuses on the similarity matching of silhouettes, as a cue for the content-based retrieval of visual information. Thus, the problem we consider is that of retrieving similar silhouettes, rather than that of recognizing objects. Silhouettes can be obtained in many different ways from images or image sequences. In obtaining experimental results, silhouettes were assumed to be complete and to have been fully extracted from images. Nevertheless, all experimental results have been obtained and evaluated in the presence of image noise, which has been modelled as noise that affects the coordinates of points on the silhouettes but not noise due self or partial occlusion. Although discussed, treatment of occlusion is left for future work.

The work presented in this paper is concerned with providing a strategy for the similarity matching and retrieval of silhouettes that is relevant to visual perception. In order to

* Corresponding author. Tel.: +30-2810-39-1600; fax: +30-2810-39-1601.

E-mail address: orphanou@ics.forth.gr (S.C. Orphanoudakis).

¹ At the time of this research, the authors were with the Institute of Computer Science, FORTH.

achieve this goal, it is important that a similarity-based silhouette retrieval method is tractable and comprehensible by end-users. In particular, the query formulation and query refinement processes should be more intuitive for these users. For these reasons, we have adopted a local representation of silhouettes based on an arc-length parameterization of its boundary. This type of description appears to be more comprehensible to observers, more compatible with visual perception than a global representation of a silhouette, and in addition, supports the partial matching of silhouette boundaries. In the remainder of this section, an overview of the proposed methodology will be presented along with relevant notation used throughout this paper.

A silhouette is a binary image produced by a planar geometric projection of a solid object. The silhouette boundary (SB) is a non-intersecting, closed 2D curve which can completely represent the silhouette.

It is possible to define signature functions on planar curves that uniquely determine the curve up to some group of transformations [2,3]. The curvature function uniquely determines a planar curve up to a Euclidean transformation and is defined as follows. Given a planar curve $C : \mathbf{R} \rightarrow \mathbf{R}^2$ as a vector function $C(\tau) = [x(\tau), y(\tau)]^T$, where T is the transpose operator and τ the Euclidean arc-length parameter, the curvature is calculated as $K(\tau) = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{3/2}}$ where $x' = \frac{\partial x(\tau)}{\partial \tau}$ [4]. The curvature function will have extrema, \mathcal{E} s, which correspond to cusps of intrusions or protrusions of the curve. Positive maxima, \mathcal{E}_+ s, and negative minima, \mathcal{E}_- s, correspond to convex and concave cusps, respectively. Inflection points, \mathcal{Z} s, where the curvature function is equal to zero, correspond to loci that the curve is locally straight. The extrema of the absolute value of the curvature function will be denoted as $|\mathcal{E}|$ s; thus $|\mathcal{E}|$ s are either \mathcal{E} s or \mathcal{Z} s. Finally, $\Theta(\sigma)$ is used to denote a source of independent and identically distributed noise with standard deviation σ .

The basic idea behind this work is the following. The conventional estimation of silhouette dissimilarity is based on the alignment and pointwise comparison of their arc-length parameterized representations (*signatures*). This comparison is performed after a scaling of the shortest signature so that both signatures exhibit the same length. It is argued that the dissimilarity estimation, obtained with such a linear scaling of two signatures, is not as precise as that obtained with an alignment that preserves correspondences. To achieve such an alignment, landmarks are used to repeatedly anchor the arc-length parameterized descriptions of the two silhouettes. Different types of landmarks are evaluated using two criteria. First, landmark localization is evaluated with respect to its robustness in the presence of noise and how this affects silhouette alignment. Second, landmark types are evaluated with respect to which one results in retrieval results that are better matched to end-user expectations. The landmark types that are evaluated are the \mathcal{E} s and \mathcal{Z} s, because they can both

be used for the selection of intrinsically defined landmarks on a silhouette. It is concluded that \mathcal{E} s yield better results, with respect to both of the above criteria.

A conventional and a novel pointwise dissimilarity metric for silhouettes are also considered and used for the retrieval of silhouettes from a large database.² It is shown that the proposed alignment method and the novel dissimilarity metric improve the precision of the retrieval results compared to traditional approaches. Both dissimilarity metrics used are generic enough and representative of two broader categories of dissimilarity metrics. This allows us to discuss the contribution of the proposed alignment method, in the similarity-based retrieval of silhouettes, independently of the dissimilarity metric used.

The remainder of this paper is organized as follows: Related work is discussed in Section 2. Robustness issues arising in the localization of landmarks are investigated in Section 3. In Section 4, a landmark-based silhouette alignment method is presented and its robustness to noise is evaluated for different types of landmarks. In Section 5, the proposed alignment method is employed in the similarity-based retrieval of silhouettes and a novel dissimilarity metric is introduced. Ideas for future work are discussed in Section 6. Section 7 is a summary.

2. Related work

There is an abundance of Cognitive and Computer Science literature on content-based image retrieval and similarity-based matching of boundaries. This section is focused on previous work that is more closely related to the approach presented in this paper. Two reviews of silhouette matching can be found in Refs. [5,6].

2.1. Cognitive studies

The process of SB perception and recognition by humans is performed in a coarse-to-fine, hierarchical manner [7,8]. That is, the larger structural features of a silhouette's shape tend to be more characteristic or salient than its details. Due to this salience, it is henceforth assumed that the size of some structure on the SB is perceptually significant in the identification or discrimination of silhouettes within a data set. Furthermore, a coarse-to-fine hierarchical representation of shape is relevant to perception [9]. Regarding silhouettes, this perceptual behavior has also been adopted in the field of Computer Vision and has been described as "scale-tuned" representation [10].

In Ref. [11], Attneave suggested that not all segments of a silhouette provide the human visual system (HVS) with an equal amount of information and that the HVS assigns more

² For many of the experiments presented in this paper the SQUID database has been used. More information can be found at <http://www.ee.surrey.ac.uk/Research/VSSP/imagdb/demo.html>.

saliency to the cusps of a silhouette. In addition to size, it has been observed that this saliency also derives from the sharpness of a cusp [12]. The sharpness of cusps has also been exploited in attempts to quantify the saliency of SB segments, for the recognition and similarity-based retrieval of SBs [10,13]. Recently, the above observations, regarding the increased saliency of cusps relative to inflection points and the correspondence of this saliency to the sharpness of a cusp, have been verified through psychophysical experiments [14–16].

Cusps have been known to play a key role in the perceptual segmentation of silhouettes into parts [17]. It has also been shown that their role in the recognition of silhouettes is important [12]; a result that has been further confirmed by measuring the performance of humans in a silhouette recognition task [18]. The “minima-rule” [17] provides a functional explanation for the majority of SBs, but there exist cases (“elbows”) in which this rule does not hold [19] and other approaches have been used to explain perceptual parsing [10]. Thus, although that perceptual parsing has not been fully explained up to now, one can argue that a perceptually relevant decomposition of silhouettes into parts should be delimited at cusps. Reports of psychophysical experiments that support this conclusion can be found in Ref. [20].

2.2. Algorithmic approaches

2.2.1. Global matching

Global or statistical approaches, such as those based on the use of moments [21] or Fourier descriptors [22], often exhibit difficulties due to the inexplicability of the result or the inability of the methods to capture perceptually relevant information, at least with respect to natural intuition. In addition, it is difficult to modify such approaches in order to achieve partial matching of silhouettes. For these two reasons, it is even more difficult for an end-user to interact with the result of a query for the purpose of query refinement. Global approaches have been used to provide a point-to-point alignment of silhouettes [23]. A successful approach that uses local rather than global statistics to establish correspondence among silhouettes can be found in Ref. [24]. However, these local statistics do not necessarily capture perceptually significant features.

2.2.2. Signature based methods

Semi-local descriptors, e.g. local curvature, facilitate the local description of the shape of a silhouette. In such cases, an arc-length parameterized description (or signature) is computed after parameterizing the curve. Given a certain parameterization, a signature can be computed that is invariant to affine transformations of a curve [25,26,2]. Such a representation exhibits an increased demand either for coordinate accuracy or well localized landmark points, which relates to the higher order of differentials that are used. The derived signatures are typically compared by computing their pointwise differences [27–29].

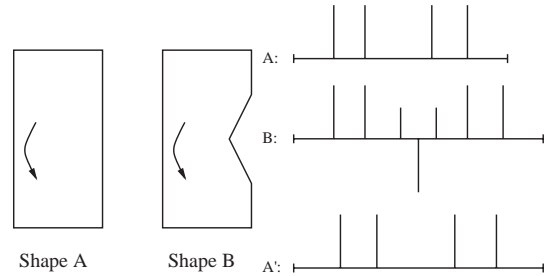


Fig. 1. Left to right: Two SBs, their corresponding, idealized, curvature functions, and the shortest one scaled.

In order for this type of comparison to be meaningful, it is required that the signature points are *aligned*. A way to deal with this problem is to establish one correspondence among the two signatures, and use this correspondence as a common origin for their comparison. This correspondence can be inferred by selecting, out of all cyclic shifts of one of the two signatures to be compared, the one that minimizes the signature difference [28]. This search yields a computational complexity of $O(N^2)$, where N is the length of the descriptions, and also requires that signatures are linearly interpolated in order to obtain the same length. However, this alignment is sensitive to noise and can also lead to counter-intuitive correspondences.

The sensitivity to noise of the above alignment can be appreciated through an example. If c_1, c_2 are two rectangles, the curvature functions would ideally exhibit 4 \mathcal{E}_+ s of equal magnitude. However, in the presence of noise, this is not necessarily the case. Similar problems are encountered in other approaches, as in selecting the common origin to be the point of the boundary that is the farthest away from its centroid [30]. In Ref. [27], the extrema of a multivalued arc-length parameterized description are selected, based on an affine-invariant parameterization of the SB.

Most importantly, aligning signatures that are based on just one correspondence typically yields counter-intuitive alignments. The linearity of the interpolation that is applied to signatures, so that they may have the same length, results in the comparison of signature segments that are not delimited by corresponding SB sites. The case is illustrated using Fig. 1: SBs A and B differ by a dent and, thus, have different arc-lengths. Let S_A and S_B be their idealized curvature-based signatures and $S_{A'}$ the linear interpolation of S_A , so that it matches the size of S_B . Supposing that a common origin has been correctly established, a simple scaling (or “stretching”) of the shortest signature causes non-corresponding signature points to be compared. In Ref. [29], a method for matching curve segments was proposed that corresponds to groupings of curve segments at a fixed resolution, without explicitly segmenting the silhouette into visual parts. More relevant to the approach presented in this paper is the scale-space matching algorithm presented in

Ref. [31]. In that work, the signature of annihilated \mathcal{L} s is compared, leading to a match that is invariant to scale, rotation, and translation. In this paper, it is argued that using \mathcal{E} s instead of \mathcal{L} s is a more robust choice.

2.2.3. Similarity measures

To date, a similarity measure for silhouettes that reflects human visual perception is still unknown. A review of perceptually relevant properties of similarity measures can be found in Ref. [32].

A conventional approach to comparing silhouettes based on their arc-length parameterized representations is the Procrustes distance [33]. Due to its generality and its utilization in the rest of this paper, this metric is presented in more detail. The Procrustes distance is formulated as follows: given two ordered sets of two-dimensional points, $P = \{p_j\}$ and $Q = \{q_j\}$, the squared Procrustes distance is given by $D^2 = \min_{\theta} \sum_j |p_j - T_{\theta} q_j|^2$, where $|\cdot|$ and T_{θ} are the length and rotation operators respectively. Invariance to the similarity and Euclidean transformations is obtained by prior normalization. By computing D , an alignment of the two silhouettes is also availed that we henceforth call *Procrustes alignment*.

Other pointwise similarity measures are the Kendall [34], the Hausdorff [5], and the one proposed in Ref. [35], but treat all points equally without considering their perceptual significance. A metric exhibiting several perceptually relevant properties is proposed in Ref. [36], but has the problem of linear scaling discussed above. Information theoretic approaches to the similarity comparison of silhouettes, such as [37], are considered in Section 5.1. Finally, elastic matching approaches can be found in Refs. [32,38], however, this paper is mainly concerned with estimating the similarity of silhouettes rather than recognizing their deformations.

2.2.4. Contributions of this work

In this paper, it is argued that the alignment method proposed in Section 4 contributes to a better utilization of existing pointwise similarity metrics. This contribution is based on the alleviation of the effect, which is due to the linear scaling of arc-length parameterized silhouette representations (see Fig. 1). The proposed approach emphasizes the need for explicitly establishing correspondences and for delimiting silhouette segments, in order to facilitate the ability of explaining the results of comparisons and to support the refinement of queries that yielded such results. In addition, the proposed alignment is robust to noise and the way the correspondences are established is relevant to visual perception.

3. Landmark selection and localization

In this section, the robustness of the localization of $|\mathcal{E}|$ s in the Mean Curvature Motion scale-space (MCM) is evaluated, leading to the conclusion that this robustness is a

monotonically increasing function of their absolute curvature value. An explanation for this phenomenon is provided and two experiments that elucidate it are presented. Initially, the scale-space analysis required for the localization of $|\mathcal{E}|$ s is reviewed.

3.1. Scale-space analysis of silhouettes for landmark selection

The pursuit of landmarks on SBs naturally leads to intrinsically defined singular points. Given the curvature-based description of a SB, points at which the curvature is minimized or maximized ($|\mathcal{E}|$) are considered as natural candidates. The MCM can be utilized to detect $|\mathcal{E}|$ s at any scale of the SB. It performs anisotropic smoothing along curves according to: $\frac{\partial C(\tau, \sigma)}{\partial \sigma} = \frac{\partial^2 C(\tau, \sigma)}{\partial \tau^2} = K(\tau) \vec{n}(\tau)$ where C is a Euclidean arc-length parameterized curve, \vec{t} and \vec{n} are its tangential and inward normal direction, and σ is the scale-parameter. For this work, a linear Gaussian smoothing of the x and y coordinate functions followed by a renormalization according to the Euclidean arc-length has been implemented [39]. Henceforth, the curvature function of a silhouette at scale n is denoted as K^n , with $n = 1$ representing the original scale.

As the curve evolves in scale, the curvature functions become simpler, since the number of extrema monotonically decreases. It can be proven [40] that, as scale increases, pairs of neighboring maxima and minima are annihilated. Coarse scale structure of a SB is more persistent under linear diffusion than fine scale structure and, thus, \mathcal{E} s occurring at a coarse scale of a SB correspond to the tips of its larger intrusions or protrusions. The scale of annihilation of an $|\mathcal{E}|$ is henceforth denoted as μ . This scale represents the *persistence* of the structure corresponding to the specific $|\mathcal{E}|$ under linear diffusion. Such a scale is of interest, due to the significance of the size of an intrusion or protrusion in the perceptually relevant identification or discrimination of silhouettes within a data set (see Section 2.1).

The detected $|\mathcal{E}|$ s can be addressed in a coarse to fine hierarchical order, given that in the linear scale-space new $|\mathcal{E}|$ s are not created as scale becomes coarser [41]. This order concerns the annihilations of $|\mathcal{E}|$ s at different scales and is obtained by selecting the first M of a descending ordering of $|\mathcal{E}|$ s with respect to their μ values. This hierarchy can be represented as a tree structure (\mathcal{T}) with SB segments as its vertices and containment relationships as its edges. At the root of the tree, one encounters the most persistent $|\mathcal{E}|$ s.

In Fig. 2, the tracking of \mathcal{E} s is demonstrated. As shown, K^1 is sensitive to noise and to the discretization of the coordinates of the SB points, which occur on a grid. As a result, spurious \mathcal{E} s occur and make cusps of the more dominant intrusions or protrusions difficult to detect. When \mathcal{E} s that occur at coarse scales are backtracked (e.g. using the algorithm in Ref. [42, Marching Cubes]), their initial locations can be retrieved.

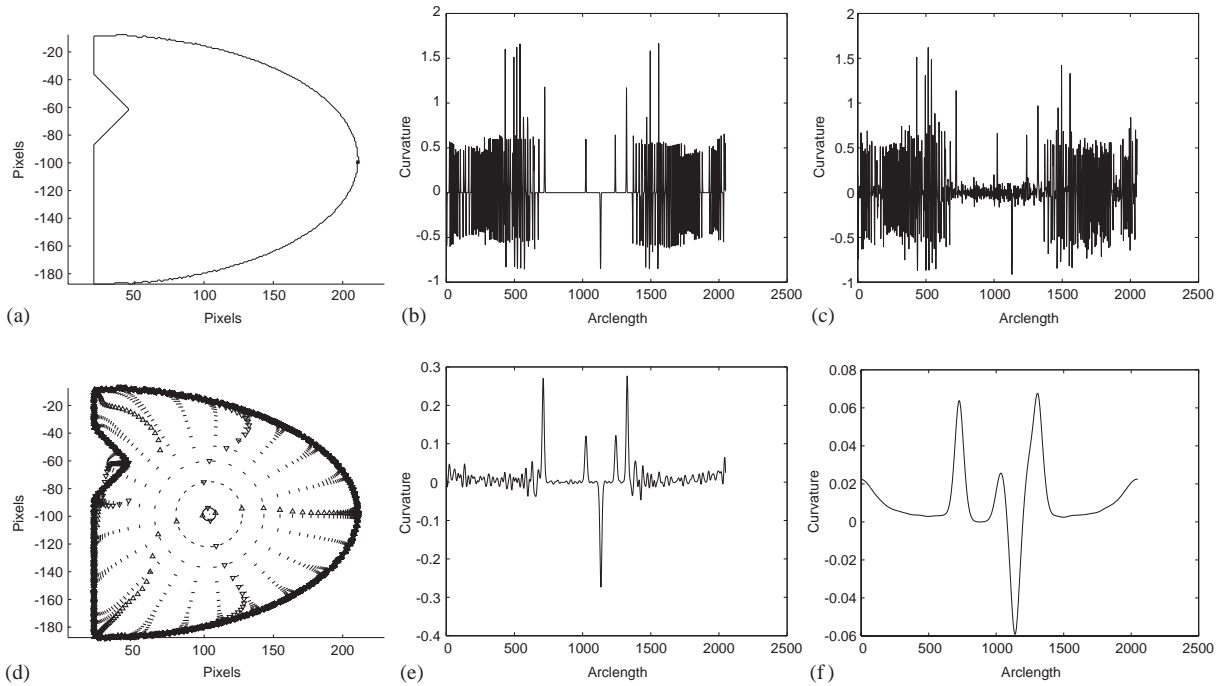


Fig. 2. (a) A SB with the origin of its coordinates at $\approx (25, 300)$. (b) The K^1 of the SB in (a): the arrangement of its points on a pixel grid creates spurious \mathcal{E} s. (c) The K^1 of the SB in (a), after application of $\Theta(0.05)$ to its coordinates, (d) Snapshots of the evolution of the noisy version of the SB in the MCM; marked are the locations of \mathcal{E} s. (e,f) K^{32} and K^{64} of the noisy version of the same SB.

3.2. Why is the localization of \mathcal{E} s more robust than that of \mathcal{L} s?

When trying to localize a \mathcal{L} , it is observed that the presence of noise can significantly alter the localization result. For example, a \mathcal{L} that occurs on a straight SB segment can be dislocated if a speckle of noise occurs on this segment: at least two \mathcal{L} s are introduced on the SB, at different loci than the locus of the original \mathcal{L} . Consequently, a backtracking process will ultimately select one of these spurious \mathcal{L} s instead of the original, which is no longer present. This effect is smaller for \mathcal{E} s, because they typically persist much longer under linear diffusion and their backtracking is initiated at a coarser scale, where their localization is more robust. This increased robustness is due to the fact that the effects of noise at coarse scales are alleviated.

The above arguments can be verified using Fig. 2. It is observed that the sharp SB cusps are still easily detected at K^{64} and can be robustly backtracked up to the 1st SB scale. By comparison, \mathcal{L} s are cluttered already at K^{32} , when they become detectable, and the probability of a mismatch is higher. In fact, the coarser the scale of a mismatch the larger the localization error, because the estimated locus of a \mathcal{L} is then determined to occur at a different SB neighborhood than the true one. The reason is that \mathcal{E} s and \mathcal{L} s are annihilated in a tree pattern (see Section 3.1) and, thus, the scale and locus

of a mismatch correspond to a tree node. When a mismatch occurs, the result is destined to occur at a different branch of the tree, rather than where the correct solution resides. The mismatch node (V) is the root of this branch as well as the root of the branch where the correct locus occurs. The closer to the root of the tree V is, the farther away the backtracking result from the pursued location will occur, since tree branches do not share nodes.

3.3. Empirical evaluation of localization robustness

In this subsection, two experiments are presented whose purpose is to investigate the robustness of the localization of potential SB landmarks, as a function of their curvature. In the first experiment, the coordinates of SBs were corrupted by noise. Then, the locations of these \mathcal{E} s were estimated and the results were compared with the true locations. The curvature of these \mathcal{E} s spanned across a wide range of values, which tended to zero at one of its sides. Thus the curve shape at the loci of \mathcal{E} s for which curvature was almost zero resembled, in practice, the shape encountered at loci of \mathcal{L} s. Furthermore, because the persistence of \mathcal{E} s (and \mathcal{L} s) is a factor that influences the level of localization robustness, the persistence of all \mathcal{E} s was set to be equal in this experiment. This control was possible by using ovals for the SBs because, ideally, they exhibit the same number of \mathcal{E} s ($=4$).

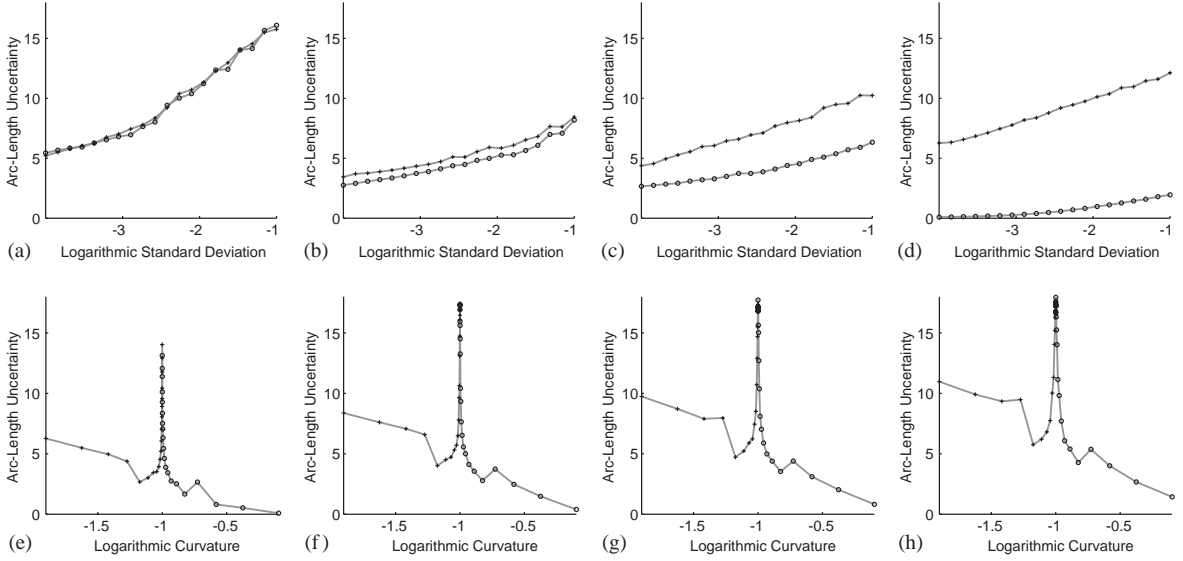


Fig. 3. Standard deviation of persistent extremum localization plotted as a function of noise level (a–d) and curvature (e–h). Circles correspond to \mathcal{E} s of relatively higher curvature and crosses to \mathcal{E} s of lower curvature. The graphs in the top row correspond to ellipses with eccentricities (a) 1.0078, (b) 1.0546, (c) 1.2336, and (d) 2. The graphs in the bottom row correspond to noise levels (values of σ) (e) 0.0001, (f) 0.0018, (g) 0.0078, and (h) 0.0336. Horizontal axes are in log-scale.

In the second experiment, simple SBs were employed for a comparison of the robustness of the localization of \mathcal{E} s against \mathcal{L} s, in order to generically demonstrate the increased robustness of the localization of \mathcal{E} s. In this experiment, the true locations of these points were unknown.

In the first experiment, the set of ellipses used featured a variety of curvature values. This was achieved by varying the eccentricity of these ellipses. The persistence of their 4 \mathcal{E} s is equal, since (a) they are annihilated in pairs and (b) ellipses are symmetrical. In the experiment, the ellipses were formulated as $[x_e(\alpha) \ y_e(\alpha)]^T = 10[e^{-1} \cos(\alpha) \ e \sin(\alpha)]^T$, where $\alpha \in [0, 2\pi]$ and e represents eccentricity. Eccentricity was exponentially increased from 1 to 2 in 20 steps. The \mathcal{E} s occur at eccentric angles $\alpha^* = \{0, \pi/2, \pi, 3\pi/4\}$ and their curvature values are $\kappa(\alpha^*) \in \{(10e^3)^{-1}, e^3/10\}$. Thus, $K(\alpha^*) \rightarrow 0.1$ for $e \rightarrow 1$ and $K(\alpha^*) \in \{0, \infty\}$ for $e \rightarrow \infty$. The noise was applied perpendicularly to the curve, as: $[x_e(\alpha) \ y_e(\alpha)]^T + \Theta(\sigma)[-y'_e(\alpha) \ x'_e(\alpha)]^T$, where prime denotes differentiation. The standard deviation σ of noise $\Theta(\sigma)$ was exponentially increased from 10^{-4} to 0.1 pixels, in 20 steps. The \mathcal{E} s were then backtracked in the MCM, applying a new pseudo-randomized Θ each time.

In Fig. 3, the standard deviation of the localization error is plotted as a function of noise level and curvature. Crosses represent \mathcal{E} s located on the “flat” parts, whereas circles represent \mathcal{E} s located on the “sharp” parts of the ellipse. Each data point is based on 10^3 localizations. In all graphs, uncertainty is mapped onto the vertical axis and measured in pixels. In the top row, σ is mapped onto the horizontal axis. The graphs correspond to ellipses of increasing eccentricity;

from a circle (leftmost) to a sharp oval (rightmost). Naturally, uncertainty increases monotonically as a function of σ . The main observation is that the localization of highly curved \mathcal{E} s is systematically less sensitive to noise than the localization of less curved \mathcal{E} s. In the leftmost graph, the two uncertainties are equal. The reason is that the ellipse was a circle and, ideally, there were no \mathcal{E} s. Thus the 4 most persistent, but spurious, \mathcal{E} s were detected as the pursued ones. In the bottom row, curvature is mapped to the horizontal axis. In the graphs, uncertainty is a monotonically decreasing function of curvature, with the exception of the peak at 0.1. At this value, the noise-free oval was a perfect circle, and since a perfect circle has a constant curvature function, there exist no real \mathcal{E} s to be detected. The detected \mathcal{E} s are due to noise and spurious by construction. The level of uncertainty at this peak is bounded by the circumference of the circle.

In the second experiment, the $M = 10$ most persistent \mathcal{E} s and the M most persistent \mathcal{L} s of some simple SBs were backtracked 500 times, after applying $\Theta(\sigma = 0.2)$ to their coordinates. Beforehand, the loci of these \mathcal{E} s were estimated in the same way, but without the presence of noise. In Fig. 4, the results of this experiment are shown with circles marking the loci detected without the presence of noise and dark points marking the detected loci of the landmarks, when noise was added. The mean errors were computed, assuming that the noise free localizations were correct, and are given in the figure caption. In this figure, it can be observed that, almost always, the noisy localizations of \mathcal{E} s and \mathcal{L} s cluster around the noise-free correct localizations. The spread of

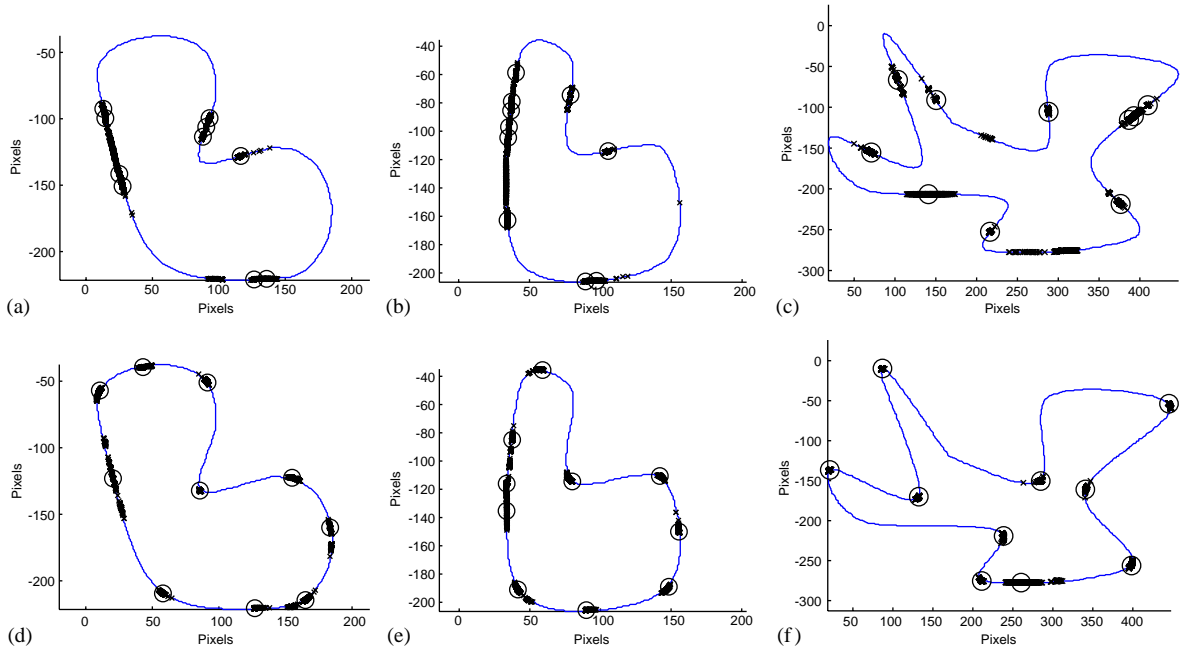


Fig. 4. Results of the localization procedure for three simple SBs, showing that this procedure is more robust for \mathcal{E} s than for \mathcal{L} s. In (a), (b), and (c) the results for the backtracking of \mathcal{L} s are shown, and in (d), (e), and (f) results are shown for the backtracking of \mathcal{E} s. Circles show the result of the noise-free localizations of these points and intonated points mark the same localizations in the presence of noise. When noise is present, the localization of \mathcal{L} s can be so erroneous that the result occurs at the locus of a different \mathcal{L} and, thus, some crosses in the right column cluster at regions where no circle is present. The mean errors were: (a) 9.409, (b) 48.207, (c) 19.33, (d) 2.006, (e) 2.114, and (f) 1.67 pixels.

these clusters is broader at regions of relatively lower curvature. In addition, for the localization of \mathcal{L} s, it is sometimes the case that noise causes the localization to occur at a locus far from the correct one. Such cases are shown in the bottom regions of graphs (a) and (c). This is an effect caused by a mismatch of \mathcal{L} s at a coarse scale in the backtracking process, as described in Section 3.2.

It is concluded that when using an arc-length parameterized SB representation and the MCM, the localization of \mathcal{E} s is more robust to noise than that of \mathcal{L} s. Thus, all other factors being equal, the structure of a SB can be landmarked with greater robustness if high-curvature landmarks are preferred over low curvature ones. It is argued that this conclusion also holds for \mathcal{E} s of the curvature function of invariant arc-length parameterizations (e.g. [2,26]). In such parameterizations, sampling density is a monotonically increasing function of curvature and, thus, the signal-to-noise ratio in this neighborhood is greater or equal.

4. Landmark-based SB alignment

In this section, a method for the alignment of a pair of SBs is proposed. This method is based on the establishment of correspondences between landmarks on these SBs fol-

lowed by the piecewise, linear interpolation of SB segments contained between corresponding landmarks. The number of these landmarks is automatically determined. The resulting alignment of SBs yields more corresponding points than alignments of SBs that are obtained by linearly stretching their arc-length parameterized descriptions. When the most persistent \mathcal{E} s are used as landmarks, the performance of the method is optimized with respect to the correctness of correspondences and robustness to noise.

Since we treat SBs without knowledge of semantics about the object that the SB portrays, the correctness of correspondences is empirically judged. We used the following criteria to evaluate this correctness: (a) similar segments that occur in the same order on two SBs should correspond, and (b) the corresponding points of these segments should occur on similar local structures.

The proposed alignment method is defined for two SBs and M landmarks on each as follows: the M cyclic shifts of the landmarks of the 2nd SB are considered. For each shift, the Procrustes distance of the shifted landmarks of the 2nd SB to the original landmarks of the 1st SB is computed. The shift that minimizes this distance is selected and the particular landmark correspondences are established. The SB segments contained between corresponding landmarks are then also considered as corresponding and they are

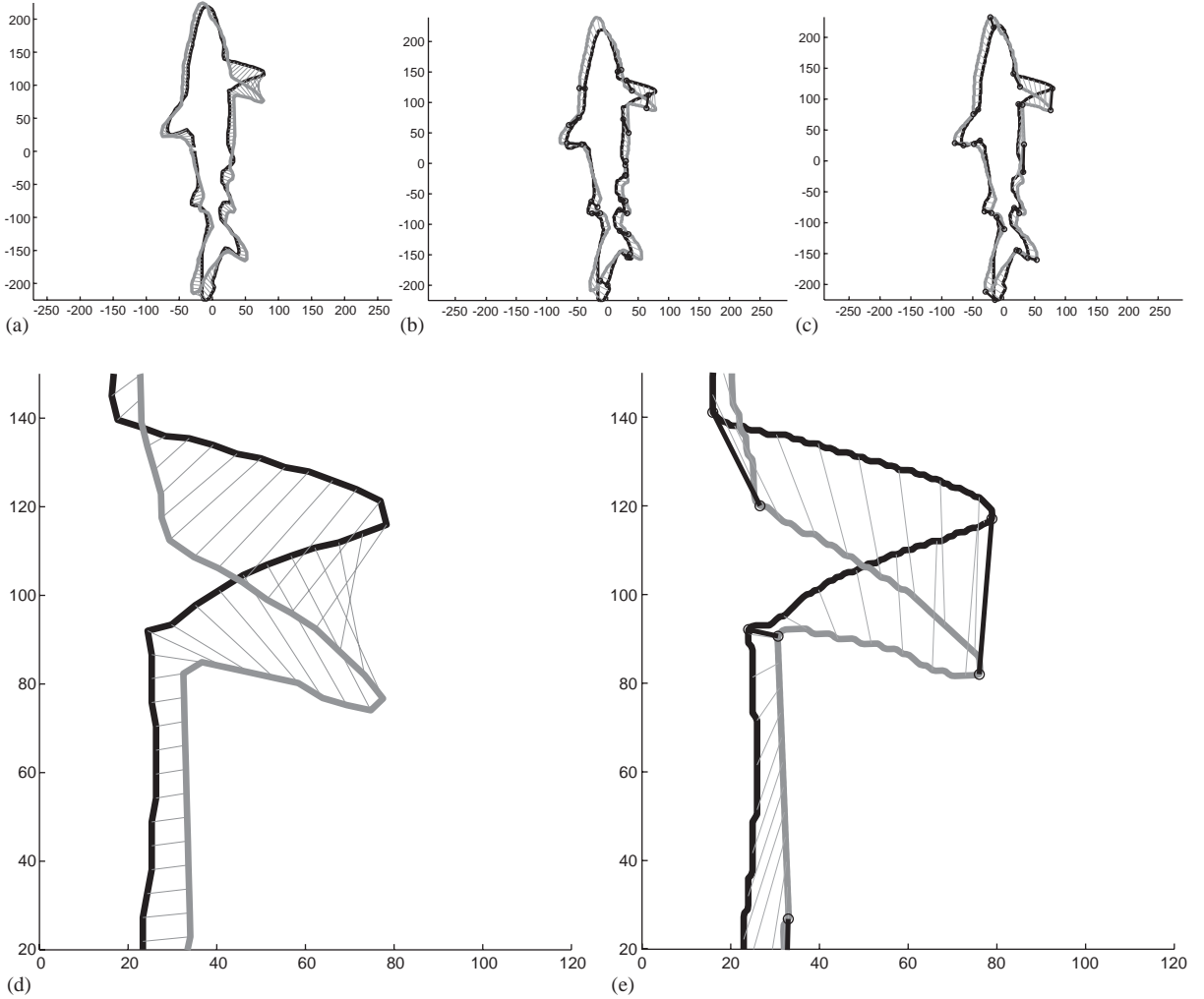


Fig. 5. Aligning in 3 different ways: (a) Procrustes-alignment, (b) most persistent \mathcal{S} s, and (c) most persistent \mathcal{E} s. Graph (d) shows a zoom into $x \in [0, 120]$, $y \in [20, 150]$ of (a), and (e) shows a zoom into the same region of (c). Black line segments mark established correspondences.

interpolated using the same number of points for each. A one-to-one correspondence is then established between these interpolation points. The resulting arc-length parameterization is piecewise linear.

The value of M is derived by computing the alignments for $M = 4, 5, \dots$ and selecting the value of M that minimizes the Procrustes distance of these M landmark points. Note that we have chosen not to normalize this distance with the cardinality of the set of landmarks (M), since this normalization was observed to result in over-fitting of the SBs. This choice, in combination with the persistent nature of the landmarks, retains the value of M at a low level. This level typically corresponds to the number of the largest structural features of the two SBs and, in our experiments, was observed to vary within the range of 3–20.

For the selection of landmarks, the most persistent $|\mathcal{E}|$ s were evaluated. This prioritization of persistent landmarks yields landmarks that correspond to the larger intrusions and protrusions of the SBs to be compared. Three advantages of this choice are the following. First, the number (M) of them that is required to align the SBs is small, which results in a reduced computational complexity due to the small number of shifts that are considered. Second, the shape of the SB region that corresponds to a scale-persistent $|\mathcal{E}|$ is less sensitive to noise due to its relatively larger size and, thus, the landmark localization process is expected to be more robust. Third, prioritization based on larger structures is perceptually more meaningful (see Section 2.1).

In Figs. 5 and 6, this alignment method is demonstrated and its performance compared against the conventional

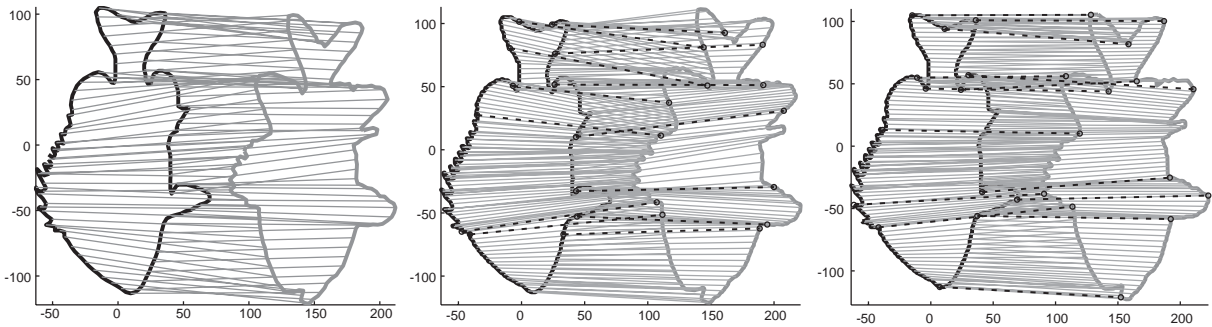


Fig. 6. Aligning in 3 different ways. Conditions as in Fig. 5: Dashed line segments mark established correspondences.

Procrustes-based SB alignment method, for which 200 sample points per SB were used. The figures demonstrate that, as a result of this piecewise alignment, there are more corresponding points in the landmark-aligned SBs than the Procrustes aligned SBs. Comparing the efficacy of \mathcal{E} s and \mathcal{L} s as landmarks for SB alignment on numerous SBs, it was observed that a better alignment is achieved using \mathcal{E} s for two reasons: (1) scale-persistent \mathcal{E} s are better SB descriptors than scale-persistent \mathcal{L} s, at least for the number of landmarks that tend to be selected by the above method and (2) the \mathcal{E} -based SB alignment exhibits increased robustness to noise. These two statements are discussed in the next two paragraphs, respectively.

4.1. Quality of descriptors

An experiment was performed to investigate the quality of description of a SB, obtained by either scale-persistent \mathcal{E} s or \mathcal{L} s. The landmarks were first detected and then used to reconstruct the contour, using cubic splines. Original and reconstructed SBs shared the same number of points. The SBs were then compared by computing their Procrustes distance. In this case, the correspondence of points between the SBs is known by construction and there is no need for alignment. Fig. 7 shows the response of the Procrustes measure for 3 reconstructions, as a function of the number of landmarks. For comparison, these plots include the response when using landmarks that are obtained through uniform sampling of the SB. The figure shows that the \mathcal{E} -based reconstruction is more similar to the original shape than the \mathcal{L} -based one. This is typically the result for the range of 3–20 landmarks that are commonly required for this alignment. The reconstruction based on uniform sampling almost always yields a smaller distance, but it is not preferred for alignment purposes, as it yields less correspondences, when point correspondence is unknown. The fluctuations in the responses for $|\mathcal{E}|$ s are due to alignment mismatches at fine scale.

4.2. Robustness of alignment

The increased robustness of the localization of \mathcal{E} s over \mathcal{L} s is observed to be inherited by the corresponding \mathcal{E} -based

alignment. In an experiment, pairs of SBs were aligned after corrupting their point coordinates with different levels of noise Θ . Fig. 8 shows the computed distance as a function of σ , using the proposed alignment method. For each noise level, the alignment was performed 100 times and the results were averaged. For comparison, the experiment was also performed for the condition where the SBs were aligned using the conventional Procrustes approach, in which 200 samples of the SBs were used. The figure demonstrates that the \mathcal{L} -based alignment is the most sensitive to noise, because dissimilarity tends to increase more rapidly as a function of noise level. Note that the increased noise sensitivity of the localization process of \mathcal{L} s can lead to a particular mis-correspondence of landmarks, which results in an erroneous alignment of the SBs. In Fig. 8(c), this can be observed by the non-monotonic behavior of the \mathcal{L} -based distance: depending on the noise level, very different alignments are established. This behavior is a result of the coarse-scale landmark mismatch in the localization of \mathcal{L} s, described in the Section 3.

5. Similarity matching of SBs

In this section, the proposed alignment method is employed in the retrieval of SBs and compared to the linear alignment of SBs, in terms of precision of the retrieval results. The purpose of this comparison is to demonstrate that the proposed alignment method increases the precision of SB dissimilarity estimation, when this dissimilarity is obtained as an accumulation of local dissimilarities along arc-length. Furthermore, a Minimum Description Length (MDL) dissimilarity measure for piecewisely aligned SBs is proposed and claimed to yield a more precise estimation of the visual dissimilarity of SBs than the Procrustes metric, thus leading to more precise results in the retrieval of SBs. In the definition of this metric the standard terminology of information theory is used (see [43] as a reference).

To perform the above comparisons, a retrieval experiment was performed under different conditions and the results obtained were compared. Under all conditions used, the SBs

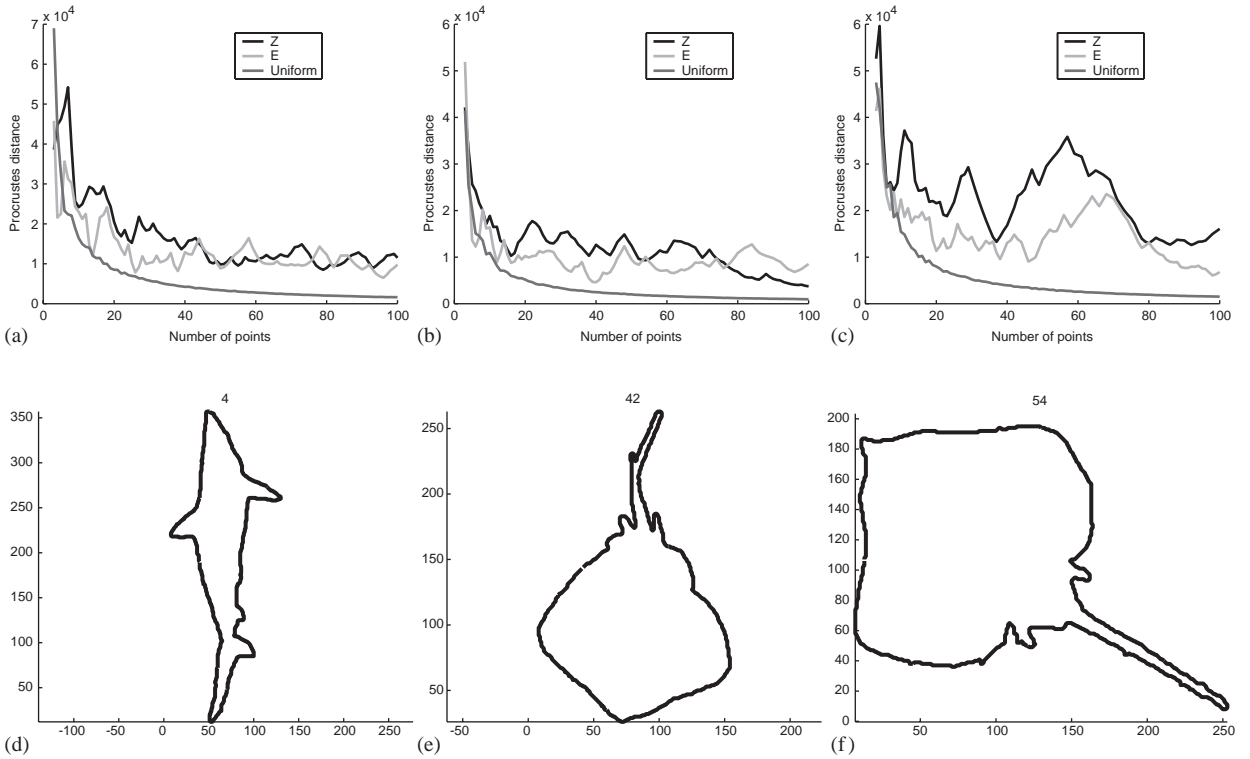


Fig. 7. Curvature extrema are better descriptors than zero crossings. Top: The graphs show the Procrustes distance between original and reconstructed silhouette, as a function of the number of points that were used for the reconstruction. When \mathcal{E} 's were used for reconstruction, the Procrustes distance is typically less than when \mathcal{Z} 's were used. The graphs of these figures were obtained for silhouettes #4, #42, and #54 (left to right) of the SQUID database. Bottom: the silhouettes used in the experiment (#4, #42, and #54).

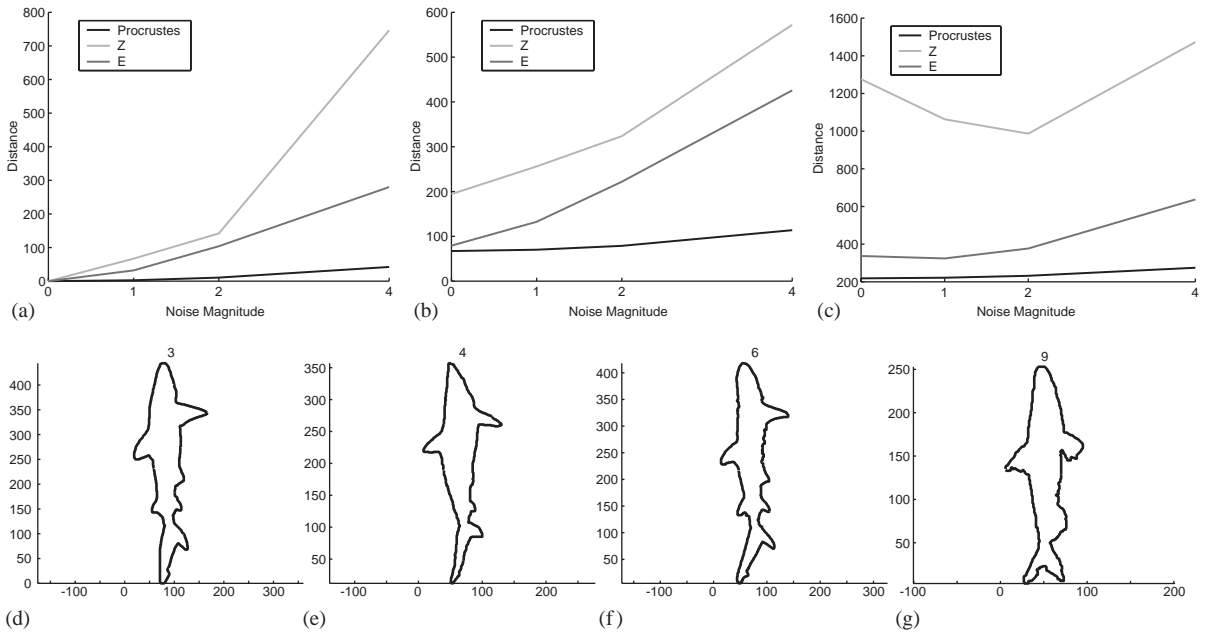


Fig. 8. Top: Estimated dissimilarity of SBs, after \mathcal{Z} -, \mathcal{E} -, and Procrustes-based alignment, as a function of noise level. The \mathcal{Z} -based alignment is observed to be the most sensitive to the increment of noise. The graphs of these figures were obtained for silhouette pairs (#4, #4), (#3, #6), and (#4, #9) (left to right) of the SQUID database. Bottom: the silhouettes used in the experiment (#3, #4, #6, and #9).

were first aligned and then their dissimilarity was estimated. To compare alignment methods we used the same similarity metric and varied the alignment method. Similarly, to comparatively evaluate dissimilarity metrics we used the same alignment method and varied the metric. Note that when performing such comparisons, neither ground truth nor results from many observers are available. Also, depending on the purpose of the visual search, one can be interested in different aspects of visual similarity: e.g. overall structural similarity, particular similarity of some detail, semantic similarity after recognition of the SB, etc. For these reasons, when comparing results from different methods, we assume that overall structural similarity and structural similarity of SB segments, whose correspondence has been established, are prerequisites to overall visual similarity.

5.1. Description of the MDL metric

The Bayesian approach to model selection is based on the Maximum A Posteriori (MAP) criterion [44],

$$m' = \arg \max_m \frac{P(d|m)P(m)}{P(d)}, \quad (1)$$

where $m \in M$ is a model from some limited model class and d is the data. Information theory leads to an alternative, but equivalent selection mechanism, where the optimal prefix code that minimizes the mean code length must have length $l = -\log_2(P_i)$ [43]. Thus, the MDL method [45] may be obtained by modifying (1) as follows:

$$m' = \arg \min_m -\log \frac{P(d|m)P(m)}{P(d)}. \quad (2)$$

Underlying the MDL approach is the notion of loss-less coding, i.e. the original data d is to be transformed in a one-to-one manner to a set of model parameters having length $-\log P(m)/P(d)$ plus the deviation of the model from the data, having length $-\log P(d|m)$. The code consists of identifying which shape is being used from the database and how many points are used in the alignment

$$L(m) = -\log(\text{number of elements in database}) + \log^*(\text{number of confident points}).$$

The code is designed in such a way that all shapes in the database are equally probable. The number of landmarks is coded by the Universal Distribution of Integers [45], $\log^*(i) = c + \log(i) + \log(\log(i)) + \dots$, summing over all positive terms.

The deviation from the model is given by the point to point alignment. In order for the decompression (decoding) to be successful, we will need to describe how the segments of the curve in the database are to be sampled, and what the

deviations are from these sample points

$$L(d|m) = \sum_i \log^*(\text{no. of samples in piece } i) + \log^*(10\sigma) - \sum_j \log(G([X_j, Y_j]^T, [x_j, y_j]^T, \sigma)).$$

The deviations are coded as a two-dimensional Gaussian distribution,

$$G([X_j, Y_j]^T, [x_j, y_j]^T, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(X_j - x_j)^2 + (Y_j - y_j)^2}{2\sigma^2}\right).$$

The point $[x_j, y_j]^T$ is coded by the inferred point $[X_j, Y_j]^T$ from the shape in the database. The standard deviation has been coded with a precision of 0.1.

5.2. Experiments

In this subsection, an experiment is presented whose purpose is twofold. The first purpose is to determine whether the proposed piecewise alignment method contributes to obtaining better retrieval results than those obtained by a standard linear alignment, when SBs are compared pointwisely. To serve this purpose, SBs were aligned either linearly or non-linearly, and their dissimilarity was estimated with the same pointwise measure. The second purpose is to compare the Procrustes and the MDL metrics, as to the precision of results that they yield when applied to piecewisely aligned SBs. Thus, SBs were aligned with the proposed method and dissimilarity was estimated using either one of these two metrics.

Each retrieval method was tested under separate conditions and was comprised of the following two stages: (a) alignment and (b) dissimilarity estimation. For stage (a), either the Procrustes alignment or the piecewise method proposed in Section 4 was utilized. For the latter method, either \mathcal{E} s or \mathcal{L} s were used as landmarks. For the stage (b), either the Procrustes distance or the MDL functional proposed in Section 5.1 was employed; we refer to these dissimilarity measures as the 1st and 2nd metric. Each combination of alignment and dissimilarity estimation approaches was evaluated under different conditions. The conditions of the experiment were:

- A: Procrustes alignment using 200 SB points and the 1st metric,
- B: proposed alignment using most persistent \mathcal{L} s as landmarks and 1st metric,
- C: proposed alignment using most persistent \mathcal{E} s as landmarks and 1st metric,

- D: proposed alignment using most persistent \mathcal{L} s as landmarks and 2nd metric,
- E: proposed alignment using most persistent \mathcal{E} s as landmarks and 2nd metric.

The experiment was performed using all 1100 SBs from the SQUID database. In all conditions, the retrieval process was the following: given a prototype, all the 1100 SB pairs were aligned and their dissimilarity values estimated; thus the prototype was also compared to itself. The dissimilarity values were then increasingly sorted. Thus, the first match was always the prototype. For ease of evaluation of the retrieval results, the retrieved silhouettes are presented after applying a rotation on the plane to their coordinates. This rotation is the one that minimizes the sum of squared distances of aligned silhouette points.

Overall, regarding the first three conditions (A, B, and C) the most appealing results were obtained under condition C. Fig. 9 shows results obtained under these conditions, for SB #28. In A, top matches exhibited small topological distances in between points whose correspondence has been established along the total arc-length of the compared SBs, but did not necessarily exhibit structural similarity when locally inspected. For example, the 3rd and 4th matches (SBs c:#154, d: #1023 in the 1st row) exhibit less structural similarity than the succeeding ones, but yield a smaller distance due to their overall similar shape. In B and C, the metric was the same as in A but primary matches exhibited more piecewise similarities. As shown in Fig. 9, more of the retrieved SBs exhibit similar intrusions and protrusions for conditions B and C than A. The results obtained in C exhibited the greatest precision out of the three first conditions. For example, SB #23 was just the 38th match in A, whereas in C it was the 3rd (see Fig. 9(q)); it can be clearly argued that this SB is more visually similar to the prototype (#28). Notice that SB #23 is also more similar to SB #28 than all matches from 3 to 7 of method A. Admittedly though, precision under condition C was less than 1 since, for example, matches 4 and 5 are less similar to the prototype than matches 6 and 7.

In Fig. 10, the results for D, E and are shown. The results of these two conditions were similar to those of B and C, since the alignment method was the same. Again, the utilization of \mathcal{E} s as landmarks yielded better results. However, the 2nd metric was observed to yield results of greater precision than the 1st. An example is the last row of Fig. 10. In that row, the retrieval precision is greater than that of, corresponding, condition C for the 1st metric (shown in Fig. 9, bottom row). For condition C, SB #684, which is not very similar to the prototype, is sorted in the 4th place, whereas for condition E, in the 6th. The above observations have been consistently repeated numerous times for all prototypes evaluated, in our experiments. In Figs. 11 and 12, more examples are given for conditions A and E.

It is concluded that when estimating dissimilarity of two SBs, based on the pointwise comparison of their aligned arc-

length parameterized representations, the precision of corresponding retrieval results is improved, if the correctness of the alignment is increased. This conclusion is based on the fact that results obtained from conditions B and C outperformed those from condition A; the only difference between these conditions was the alignment stage. The reason is that the Procrustes-based alignment aims in the minimization of point distances, but sacrificing correctness of correspondences. Thus distances of non-corresponding points are likely to be accounted in the dissimilarity estimation.

A motivating factor to evaluate retrieval precision for different landmark types was to test if any of these types would systematically yield an intuitively better alignment and, in turn, more precise retrieval results. We systematically obtained more precise results when using the most persistent \mathcal{E} s as landmarks both for the 1st and 2nd similarity metrics, in condition pairs (B, C) and (D, E) respectively. We attribute this precision to the generically better alignment that was obtained for \mathcal{E} s (see Section 4).

Regarding the dissimilarity estimation, it was observed that the results obtained using the 2nd metric better matched the expectations of observers. In particular, there was an increase of the precision of retrieval, from condition C to condition E as well as from B to D. These two pairs of conditions differed only at the dissimilarity metric. From our results, we conclude that the MDL metric leads to more precise results than the Procrustes. A reason for this is that the 2nd metric accumulates a small penalty for small disparities of SB points whose correspondence has been established and, in contrast, penalizes large disparities more. Thus, when correctly aligned, two locally similar SB segments yield a particularly small dissimilarity value.

6. Discussion

The two main reasons that caused this work to be focused in the field of local methods for the comparison of SBs instead of global ones, were: (a) the ability to explain the retrieval results and (b) the ability of the alignment method to facilitate partial matching and/or query enhancement.

Explaining the similarity results in terms of dissimilarities along arc-length and in particular in terms of segments appears to be an intuitive way to do so, since the perception of silhouettes occurs in a parsed way (see Section 2.1). Using the proposed alignment and some dissimilarity metric, it is possible to explain the dissimilarity estimation results as the accumulation of the dissimilarities of segments instead of points. For example in Fig. 13, such a way is demonstrated by averaging all the pointwise dissimilarities of two segments whose correspondence has been established and regarding this average as the dissimilarity value of these segments. The explanation as visualized appears to be meaningful, because the alignment was fairly successful and, thus, the averages approximate the perceptually judged dissimilarities of segments for which correspondence has been established.

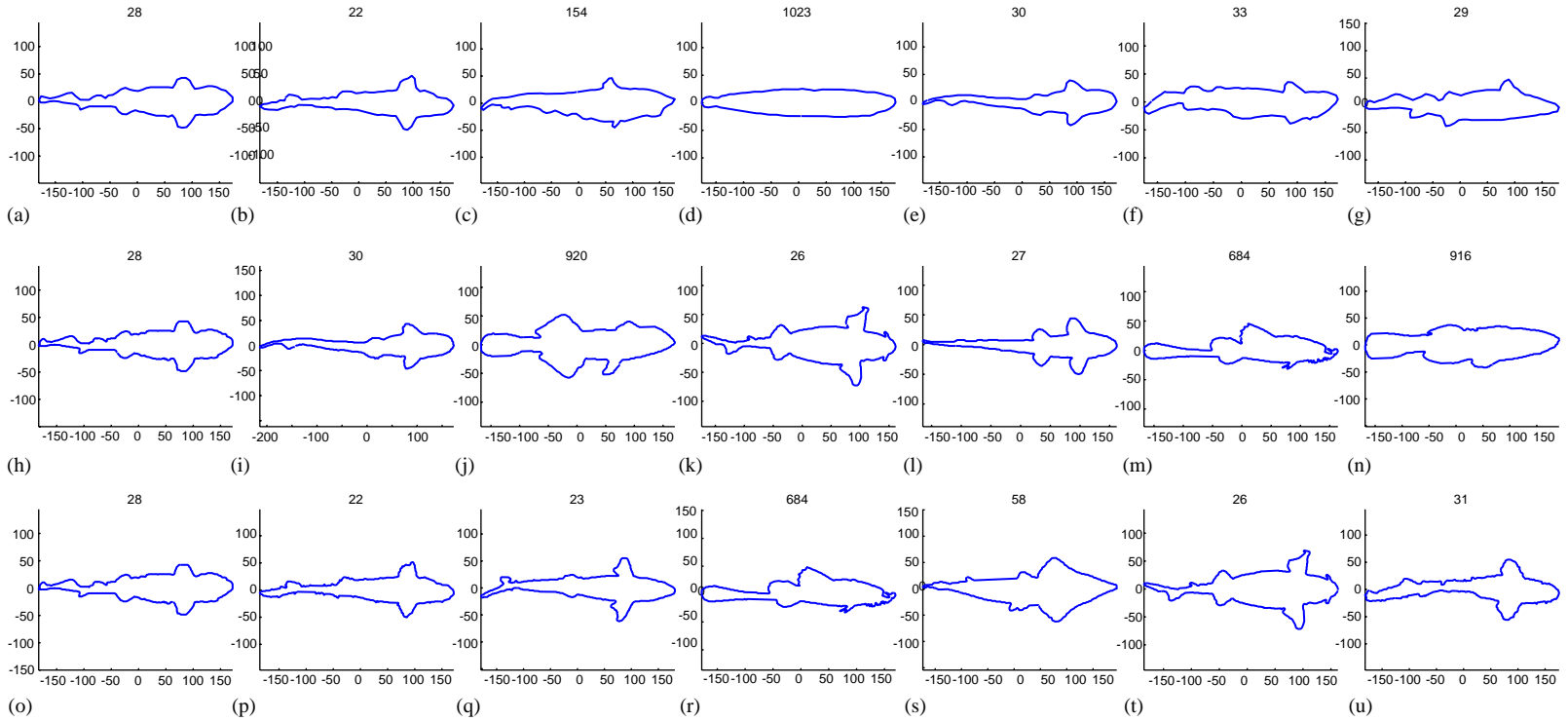


Fig. 9. Top matches of the three first conditions of the retrieval experiment presented in this section for prototype #28. Rows, from top to bottom, correspond to conditions A, B, and C, respectively. Estimated dissimilarity grows from left to right.

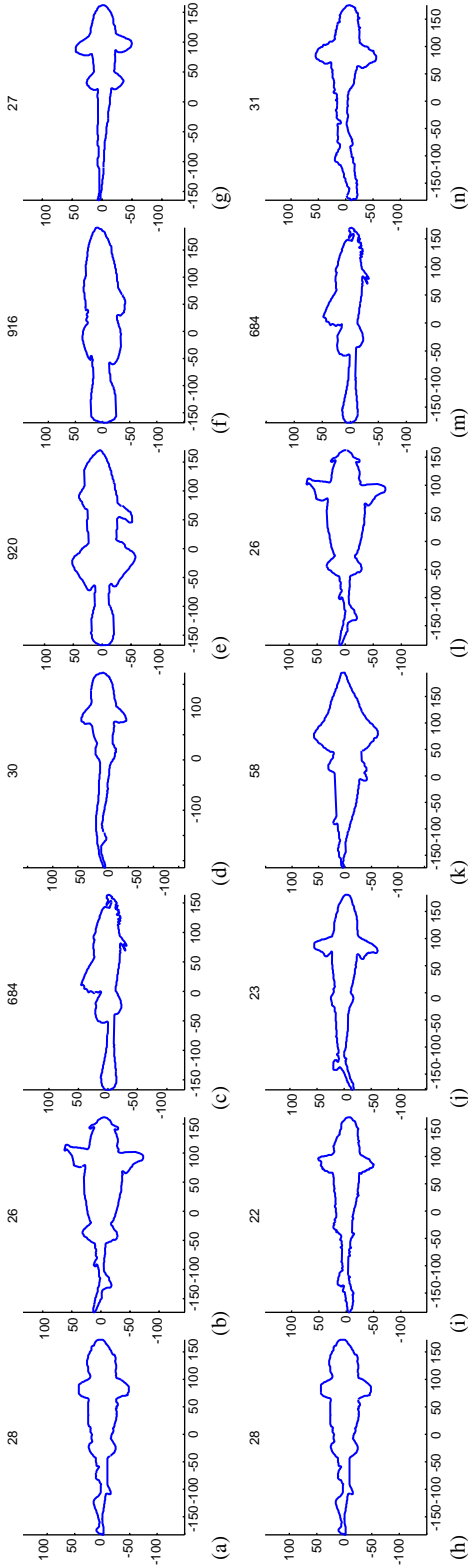


Fig. 10. Top matches of the last two conditions of the retrieval experiment presented in this section, for prototype #28. Arrangement of results as in Fig. 9.

It has to be pointed out though that the SB segments obtained by the proposed method do not correspond to parts, as defined perceptually. We regard research in understanding perceptual parsing as important for their retrieval, because adjacent segments could be assembled into parts and form predicates of a partial SB query. Thus one could search for SBs that exhibit great similarity regarding two distinct SB regions, which are segmented by an equal number of adjacent landmarks and arranged in the same order. A simpler, but less automatic, way to do so is to perform this assembly manually through a user-interface. For example, in Fig. 13 a user could select the segments corresponding to the head of the fish and search for SBs that exhibit increased similarity just for the segments that correspond to these pre-selected segments of the head. We regard that both the incorporation of cognitive knowledge in such an effort, as well as the development of an appropriate graphical user-interface, require greater attention and have postponed such efforts until a future paper. In terms of this research avenue, landmark points have been proposed (most persistent \mathcal{E} s) that are relevant to the perceptual parsing of silhouettes, as this is suggested to be mainly determined by cusps of large intrusions or protrusions (see Section 2.1).

Clearly, the ability to locally compare SBs contributes more to the facilitation of partial matching in a visually intuitive way than global methods. Nevertheless, semi-global approaches (e.g. [24]) exhibit a decreased dependence on the establishment of correspondences, but are difficult to intuitively generalize for partial matching and also cannot explain the results in an intuitive way for end-users.

The ability to find corresponding landmarks as well as segments in SBs can facilitate the computational optimization of partial matching. In particular, when SBs are compared in pointwise fashion and when one attempts to match a given SB segment to some part of a whole SB the following problem arises. The factor by which the description of the given segment has to be scaled in order to match some segment of the whole SB is unknown and, thus, one may have to adopt a computationally complex solution. E.g. optimize a partial similarity metric over numerous scalings of the given segment and different segmentations of the whole SB. Instead, if point correspondences are assumed, the evaluated scalings can be restricted to those that (approximately) align landmarks on the segment and the SB together.

7. Summary

In this paper, a method was proposed for the piecewise alignment of SBs that yields more correct correspondences than aligning the SBs by a linear scaling of the arc-length parameterization of one of them. The method is based on the correspondence establishment of landmarks and its performance is optimized when most persistent \mathcal{E} s are used, because their localization is more robust than that of other intrinsically defined points on the SB's. These landmarks are

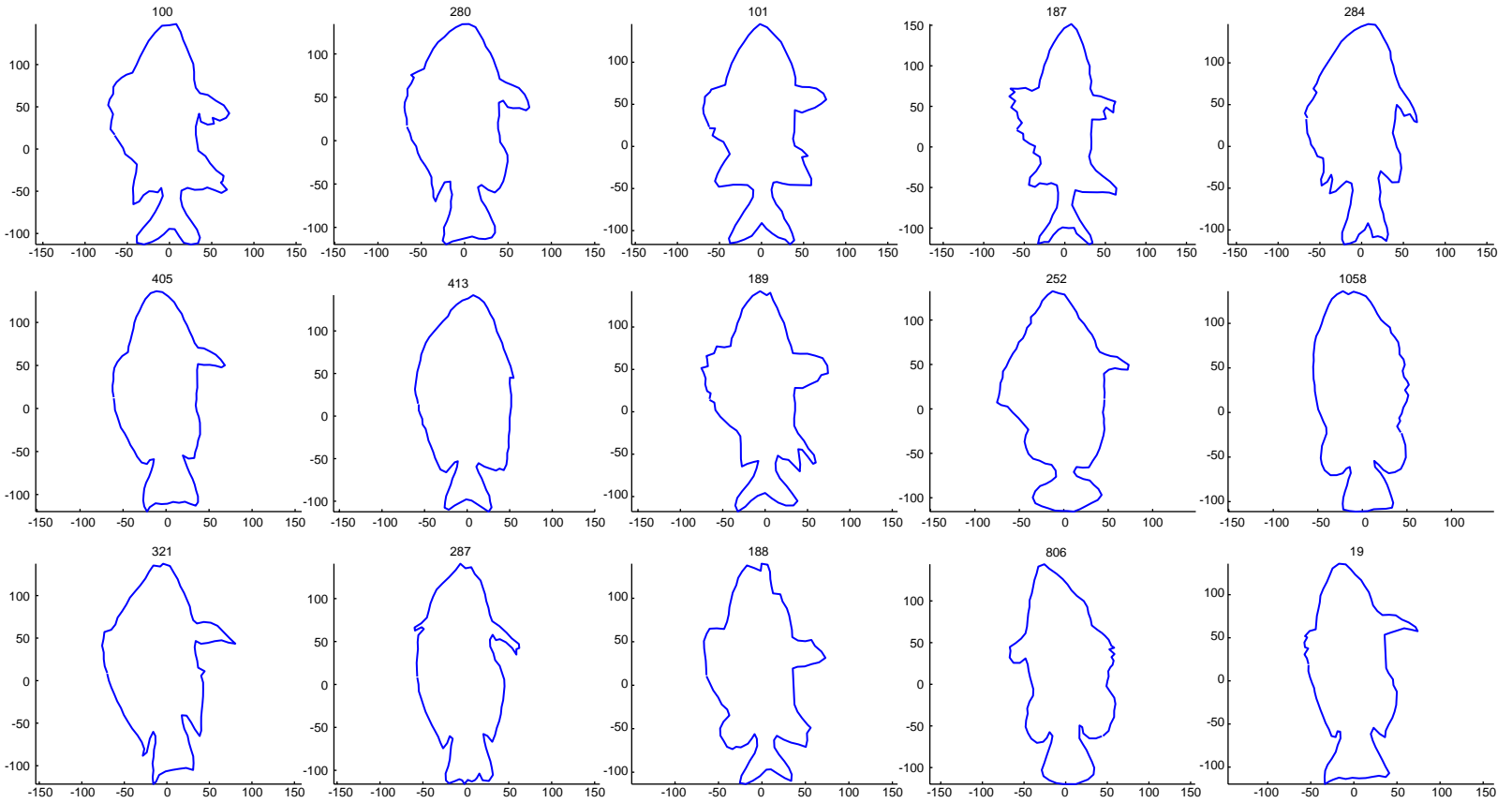


Fig. 11. Top matches of condition A of the retrieval experiment presented in this section, for prototype #100. Arrangement of results as in Fig. 9.

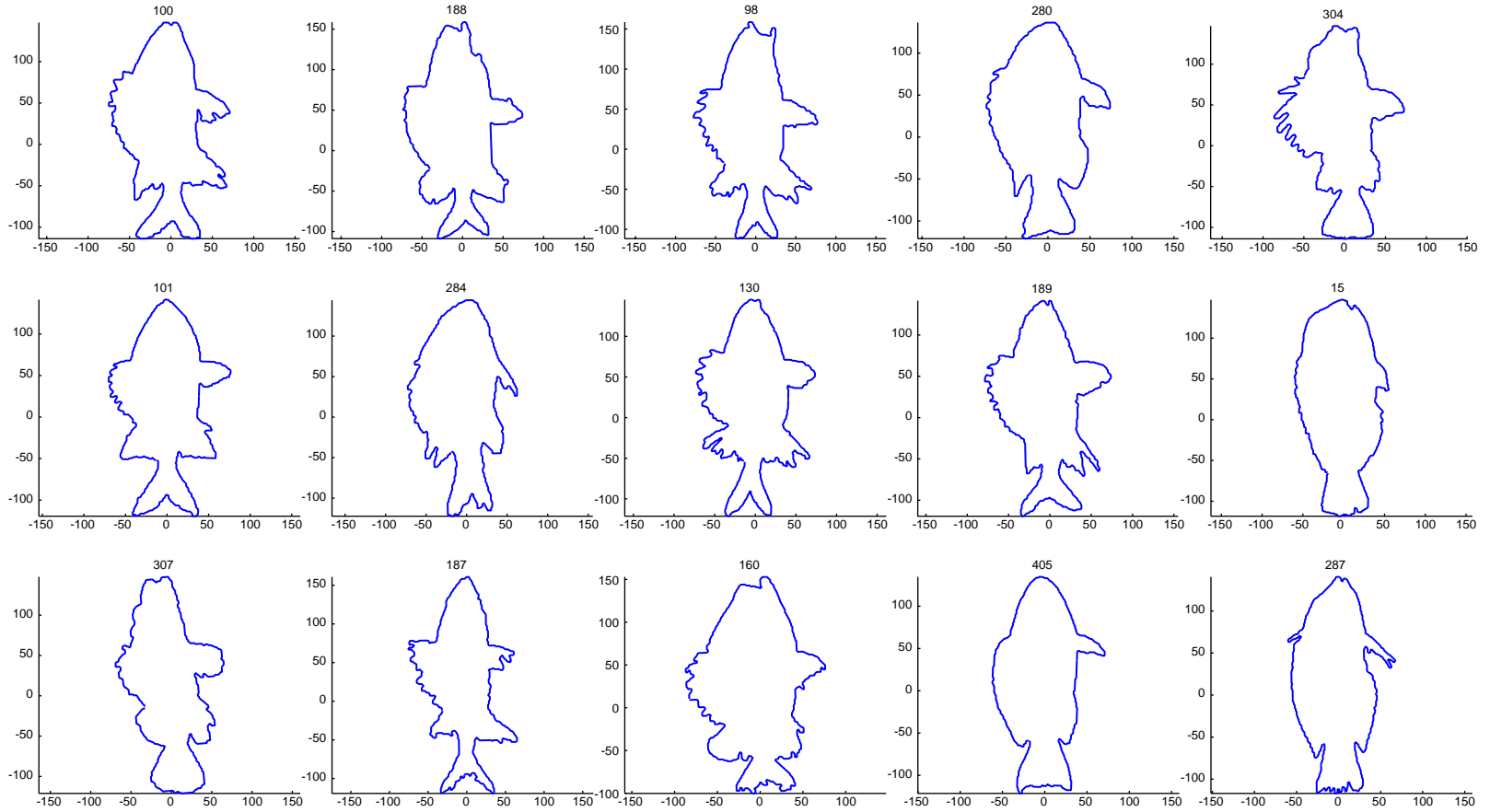


Fig. 12. Top matches of condition E of the retrieval experiment presented in this section, for prototype #100. Arrangement of results as in Fig. 9.

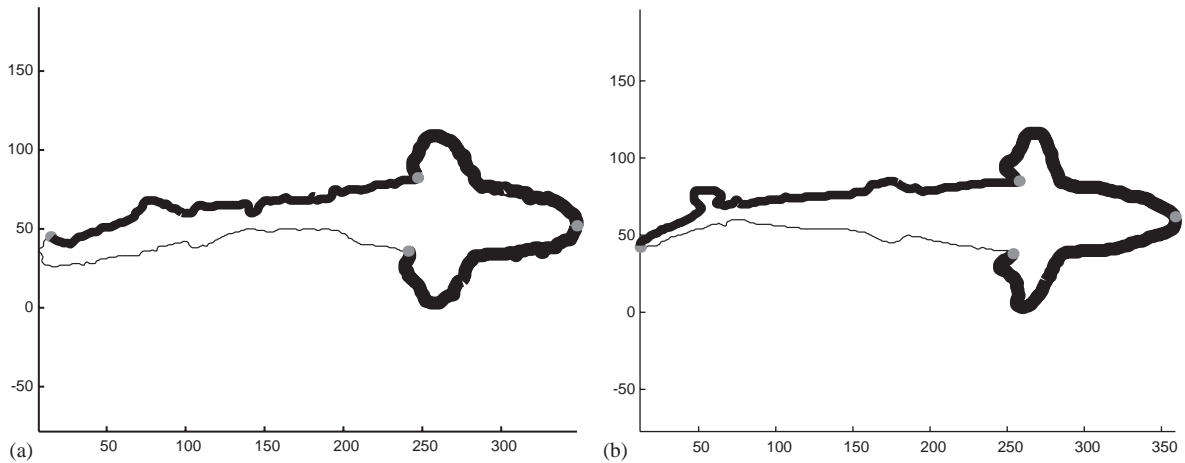


Fig. 13. Calculation of dissimilarity per segment. The width of the line represents the level of similarity between the corresponding segments of these two, aligned, SBs.

relevant to human visual perception and the results obtained are intuitively comprehensible and explainable as well as compatible to expectations of observers. It was also shown that when using this alignment method to pointwisely compare two SBs, more precise results are obtained than when performing the same comparison but after linearly aligning the SBs.

Naturally, the quantification of dissimilarity after the pointwise comparison of SBs is important to the precision of retrieval queries. The Procrustes and a MDL dissimilarity metrics were tested to show that the latter is more precise. Due to the generality of these metrics, it is asserted that the proposed alignment method could contribute to methods that use more sophisticated arc-length parameterized dissimilarity metrics.

References

- [1] Merriam-Webster. Merriam-Webster's Collegiate Dictionary, Merriam-Webster, Inc., 2003.
- [2] T. Moons, E.J. Pauwels, L.J. van Gool, A. Oosterlinck, Foundations of semi-differential invariants, *Int. J. Comput. Vision* 14 (1995) 25–47.
- [3] E.J. Pauwels, T. Moons, L.J. van Gool, P. Kempenaers, A. Oosterlinck, Recognition of planar shapes under affine distortion, *Int. J. Comput. Vision* 14 (1995) 49–65.
- [4] M.P. Do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, NJ, 1976.
- [5] R.C. Veltkamp, M. Hagedoorn, *Principles of Visual Information Retrieval*, chapter: state of the art in shape matching, *Advances in Pattern Recognition*, Springer, Berlin, 2001, pp. 87–119.
- [6] D. Zhang, G. Lu, Review of shape representation and description techniques, *Pattern Recognition* 37 (1) (2004) 1–19.
- [7] J.J. Koenderink, The structure of images, *Biol. Cybernet.* 50 (1984) 363–370.
- [8] A.P. Witkin, Scale-space filtering, *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI '83)*, vol. 2, pp. 1019–1022, Karlsruhe, Germany, August 1983.
- [9] S.E. Palmer, Hierarchical structure in perceptual representation, *Cognitive Psychol.* 9 (1977) 441–474.
- [10] K. Siddiqi, B.B. Kimia, Parts of visual form: computational aspects, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (3) (1995) 239–251.
- [11] F. Attneave, Some informational aspects of visual perception, *Psychol. Rev.* 61 (1954) 183–193.
- [12] D.D. Hoffman, M. Singh, Saliency of visual parts, *Cognition* 63 (1997) 29–78.
- [13] L.J. Latecki, R. Lakmper, Convexity rule for shape decomposition based on discrete contour evolution, *Comput. Vision Image Understand.* 73 (3) (1999) 441–454.
- [14] C. Lamote, J. Wagemans, Rapid integration of contour fragments: from simple filling-in to parts-based shape description, *Visual Cogn.* 6 (1999) 345–361.
- [15] J.F. Norman, P. Phillips, H.E. Ross, Information concentration along the boundary contours of naturally shaped solid objects, *Perception* 30 (2001) 1285–1294.
- [16] J. De Winter, S. Panis, J. Wagemans, Perceptual saliency of points along the contour of everyday objects: a large-scale study (abstract), *J. Vision* 2 (7) (2002) 487a DOI 10.1167/2.7.487.
- [17] D.D. Hoffman, W.A. Richards, Parts of recognition, *Cognition* 18 (1984) 65–96.
- [18] J. Wagemans, J. De Winter, S. Panis, The awakening of attneave's sleeping cat: identification of everyday objects on the basis of straight-line versions (abstract), *J. Vision* 2 (7) (2002) 674a DOI 10.1167/2.7.674.
- [19] M. Singh, G. Seyranian, D.D. Hoffman, Parsing silhouettes: the short-cut rule, *Perception Psychophys.* 61 (1999) 636–660.

- [20] J. De Winter, J. Wagemans, Segmentation of object outlines into parts: From a large-scale normative study to a model (abstract), *J. Vision* 1 (3) (2001) 421a DOI 10.1167/1.3.421.
- [21] A. Khotanzad, Y.H. Hong, Invariant image recognition by zernike moments, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (5) (1990) 489–497.
- [22] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, *IEEE Trans. Comput. C* (21) (1972) 269–281.
- [23] S. Sclaroff, A.P. Pentland, Modal matching for correspondence and recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (6) (1995) 545–561.
- [24] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (4) (2002) 509–522.
- [25] A. Bruckstein, R. Holt, A. Netravali, T. Richardson, Invariant signatures for planar shape recognition under partial occlusion, *CVGIP: Image Understand.* 58 (1) (1993) 49–65.
- [26] L.J. van Gool, T. Moons, E.J. Pauwels, A. Oosterlinck, Vision and Lie's approach to invariance, *Int. J. Comput. Vision* 13 (4) (1995) 259–277.
- [27] M. Klot, E. Rivlin, Invariant-based shape retrieval in pictorial databases, *Comput. Vision Image Understand. CVIU*, 71 (2) (1998) 182–197.
- [28] J.R. Parker, Z. Zhou, Object recognition using signatures, *Proceedings of the International Conference on Signal and Image Processing*, Las Vegas, Nevada - USA, 1998, pp. 27–31.
- [29] L.J. Latecki, R. Lakmper, Application of planar shape comparison to object retrieval in image databases, *Pattern Recognition* 35 (1) (2002) 15–29.
- [30] R.C. Gonzales, R.E. Woods, *Digital Image Processing*, Addison-Wesley, Reading, MA, 1992.
- [31] F. Mokhtarian, Silhouette-based isolated object recognition through curvature scale space, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (5) (1995) 539–544.
- [32] R. Basri, L. Costa, D. Geiger, D. Jacobs, Determining the similarity of deformable shapes, *Vision Res.* 38 (15–16) (1998) 2365–2385.
- [33] F.L. Bookstein, Shape and the information in medical images: a decade of morphometric synthesis, *Comput. Vision Image Understand.* 66 (2) (1997) 97–118.
- [34] D.G. Kendall, Shape-manifolds, procrustean metrics and complex projective spaces, *Bull. Lond. Math. Soc.* 16 (1984) 81–121.
- [35] P. Giannopoulos, R.C. Veltkamp, A pseudo-metric for weighted point sets, In *ECCV02*, Copenhagen, Denmark, vol. 3, 2002, pp. 715–730.
- [36] L.J. Latecki, R. Lakmper, Shape similarity measure based on correspondence of visual parts, *IEEE Trans. Pattern Anal. Mach. Intell.* 2 (10) (2000) 1185–1190.
- [37] R. Davies, T.F. Cootes, C.J. Taylor, A minimum description length approach to statistical shape modelling, *Proceedings of the 14th International Conference on Information Processing in Medical Imaging*, 2001, pp. 50–63.
- [38] A. del Bimbo, P. Pala, Visual image retrieval by elastic matching of user sketches, *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (2) (1997) 121–132.
- [39] F. Mokhtarian, A. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1) (1986) 34–43.
- [40] R. Teixeira, Curvature motions, medial axes and distance transforms, Ph.D. Thesis, Department of Mathematics, Harvard University, Cambridge, MA, USA, June 1998.
- [41] L. Florack, A. Kuijper, The topological structure of scale-space images, *J. Math. Imag. Vision* 12 (1) (2000) 65–80.
- [42] J.D. Foley, A. van Dam, S.K. Feiner, J.F. Hughes, *Computer Graphics, Principles and Practice*, 2nd Edition, Addison-Wesley, Reading, MA, 1997.
- [43] C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.
- [44] A.F.M. Smith, J.M. Bernardo, *Bayesian Theory*, Wiley, New York, 1993.
- [45] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, Singapore, 1989.

About the Author—XENOPHON ZABULIS is a Postdoctoral Fellow at the Computer and Information Science Department, University of Pennsylvania and affiliated with the interdisciplinary GRASP laboratory at the same institution. He is currently working in the field of multiple-view scene acquisition for tele-immersion. He obtained a Ph.D. in 2001, a M.A. in 1998, and a B.A. in 1996, in Computer Science from the University of Crete. Prior to his current appointment he was a Postdoctoral Fellow at the Institute for Research in Cognitive Science, University of Pennsylvania, working in the understanding of human binocular vision. His research interests include visual information retrieval by content as well as computational aspects of human and multiple-view vision.

About the Author—JON SPORRING received his Master and Ph.D. degree from the Department of Computer Science, University of Copenhagen, Denmark in 1995 and 1998, respectively. Part of his Ph.D. program was carried out at IBM Research Center, Almaden, California, USA. Following his Ph.D., he worked as a visiting researcher at the Computer Vision and Robotics Lab at Foundation for Research & Technology - Hellas, Greece, and as assistant research professor at 3D-Lab, School of Dentistry, University of Copenhagen. Since 2003 he has been employed as associate professor at the Department of Computer Science, University of Copenhagen. His main topic of research is medical image processing, medical computer graphics, and information theory.

About the Author—STELIOS C. ORPHANOUDAKIS holds a Ph.D. degree in Electrical Engineering from the Thayer School of Engineering, Dartmouth College, USA, a M.S. degree in Electrical Engineering from the Massachusetts Institute of Technology (MIT), and a B.A. degree from Dartmouth College. He held a faculty appointment in the Departments of Diagnostic Radiology and Electrical Engineering at Yale University, USA, from 1975 until 1991. Since 1986, he holds a faculty appointment as Professor of Computer Science at the University of Crete, Greece. Furthermore, from 1991 until 1994, he was Acting Director of the Institute of Computer Science, Foundation for Research and Technology-Hellas (ICS-FORTH), and from 1994 until 2004 he was Director of this Institute. Today, he is Chairman of the Board of Directors of FORTH and President of the European Research Consortium for Informatics and Mathematics (ERCIM). At ICS-FORTH, he is scientific leader of the Center for Medical Informatics and Health Telematics Applications and the Computational Vision and Robotics

Laboratory. Prof. Orphanoudakis is a member of many honorary and professional societies and a Senior Member of the Institute of Electrical and Electronics Engineers (IEEE). He has many years of academic and research experience in the fields of computational vision and robotics, intelligent image management and retrieval by content, medical informatics, and medical imaging. He is the author of more than 130 publications in international scientific journals, refereed conference proceedings and books. He has served on various committees and working groups of the European Commission and has been active in European R & D programs. During the period 1995–2000, he served on the National Telecommunications and Post Commission of Greece. Finally, from 1994 until 2001 he served on the National Advisory Research Council of Greece and, from 1998 until 2002, on the Board of Directors of the Hellenic Foundation for Culture.